

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER: \_\_\_\_\_**

**IMAGES ARE BEST AVAILABLE COPY.**

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 11-272710

(43)Date of publication of application : 08.10.1999

(51)Int.Cl.

G06F 17/30

(21)Application number : 10-092645

(71)Applicant : OMRON CORP

(22)Date of filing : 20.03.1998

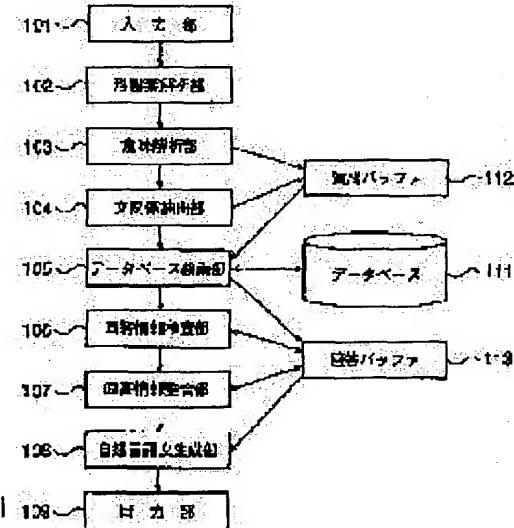
(72)Inventor : KAMO MASAMITSU  
KITAGAWA SHOJI  
YATAKE HIDENORI  
KOJIYA KAZUTO

## (54) INFORMATION RETRIEVAL SYSTEM, INFORMATION RETRIEVAL METHOD, AND STORAGE MEDIUM

### (57)Abstract:

**PROBLEM TO BE SOLVED:** To provide an information retrieval system, an information retrieval method, and a record medium which present information meeting a retrieval request so that it may be easy for a user to understand by arranging and classifying information in accordance with plural intends of the user or a complicated condition.

**SOLUTION:** Input sentences are analyzed and are decomposed to elements by a morpheme analysis part 102, and input sentences are converted to an internal expression form to clarify their concept and conceptual relations by a semantic analysis part 103, and relations between sentences of input sentences are extracted by a sentence relation extraction part 104, and a data base 111 is accessed based on the internal expression form by a data base retrieval part 105 to perform information retrieval and information registration, and duplicate parts of answer information as the result of information retrieval are extracted by an answer information check part 106, and duplicate parts of answer information are adjusted to make the presentation order consistent by an answer information adjustment part 107, and answer information adjusted and made consistent is converted to a natural language form by a natural language sentence generation part 108.



## LEGAL STATUS

[Date of request for examination] 20.06.2003

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of

rejection]

[Date of requesting appeal against examiner's  
decision of rejection]

[Date of extinction of right]

Copyright (C) 1998,2003 Japan Patent Office

(51)Int.Cl.<sup>6</sup>  
G 0 6 F 17/30

識別記号

F I  
G 0 6 F 15/403  
3 3 0 C  
3 7 0 Z

## 審査請求 未請求 請求項の数10 FD (全 12 頁)

(21)出願番号 特願平10-92645

(22)出願日 平成10年(1998)3月20日

(71)出願人 000002945  
オムロン株式会社  
京都府京都市右京区花園土堂町10番地  
(72)発明者 加茂 正充  
京都府京都市右京区花園土堂町10番地 オ  
ムロン株式会社内  
(72)発明者 北川 昇治  
京都府京都市右京区花園土堂町10番地 オ  
ムロン株式会社内  
(72)発明者 八竹 英紀  
京都府京都市右京区花園土堂町10番地 オ  
ムロン株式会社内  
(74)代理人 弁理士 飯塚 信市

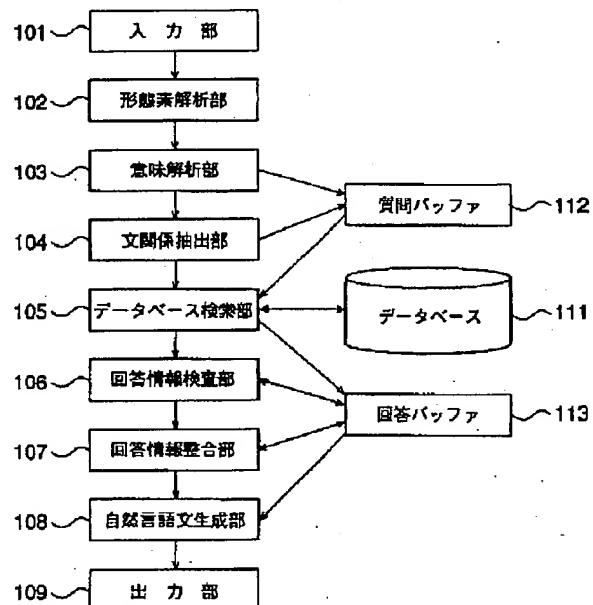
最終頁に続く

## (54)【発明の名称】 情報検索システム、情報検索方法および記録媒体

## (57)【要約】

【課題】 ユーザの複数の意図や複雑な条件に対して情報の整理・分類を行って、検索要求に即した情報をユーザーにわかりやすく提示する情報検索システム、情報検索方法および記録媒体を提供する。

【解決手段】 形態素解析部102で入力文を解析して要素に分解し、意味解析部103により入力文を内部表現形式に変換して該入力文の概念および概念関係を明確にし、文関係抽出部104により入力文における文と文の間の関係を抽出し、データベース検索部105により内部表現形式に基づきデータベース111をアクセスして情報検索および情報登録を行い、回答情報検査部106により情報検索の結果である回答情報の重複部分を抽出し、回答情報整合部107により回答情報の重複部分を調整し提示順序の整合をとり、自然言語文生成部108により調整および整合後の回答情報を自然言語形式に変換する。



## 【特許請求の範囲】

【請求項1】 ユーザからの入力文を解析して要素に分解する形態素解析手段と、前記形態素解析結果に基づき、前記入力文を内部表現形式に変換して該入力文の概念および概念関係を明確にする意味解析手段と、前記入力文における文と文の間の関係を抽出する文関係抽出手段と、前記内部表現形式に基づき、データベースをアクセスして情報検索および情報登録を行うデータベース検索手段と、前記情報検索の結果である回答情報の重複部分を抽出する回答情報検査手段と、前記回答情報の重複部分を調整し提示順序の整合をとる回答情報整合手段と、前記調整および整合後の回答情報を自然言語形式に変換する自然言語生成手段と、を具備することを特徴とする情報検索システム。

【請求項2】 前記回答情報検査手段における調整および整合は、前記意味解析手段による文概念および概念関係情報、並びに、前記文関係抽出手段による文関係情報に基づいて行われることを特徴とする請求項1に記載の情報検索システム。

【請求項3】 前記意味解析手段による前記内部表現形式、文概念および概念関係情報、並びに、前記文関係抽出手段による文関係情報を一時的に保持する質問バッファと、前記データベース検索手段による回答情報を一時的に保持する回答バッファと、を具備することを特徴とする請求項1または2に記載の情報検索システム。

【請求項4】 前記回答情報検査手段における調整および整合は、前記回答情報が複数ある場合に、ユーザが前提条件として使用した単語を元に基準を定め、分類して該回答情報を提示することを特徴とする請求項1、2または3に記載の情報検索システム。

【請求項5】 ユーザからの入力文を解析して要素に分解する形態素解析ステップと、前記形態素解析ステップの結果に基づき、前記入力文を内部表現形式に変換して該入力文の概念および概念関係を明確にする意味解析ステップと、前記入力文における文と文の間の関係を抽出する文関係抽出ステップと、前記内部表現形式に基づき、データベースをアクセスして情報検索および情報登録を行うデータベース検索ステップと、前記情報検索の結果である回答情報の重複部分を抽出する回答情報検査ステップと、前記回答情報の重複部分を調整し提示順序の整合をとる回答情報整合ステップと、前記調整および整合後の回答情報を自然言語形式に変換

する自然言語生成ステップと、を具備することを特徴とする情報検索方法。

【請求項6】 前記回答情報検査ステップにおける調整および整合は、前記意味解析ステップによる文概念および概念関係情報、並びに、前記文関係抽出ステップによる文関係情報に基づいて行われることを特徴とする請求項5に記載の情報検索方法。

【請求項7】 前記回答情報検査ステップにおける調整および整合は、前記回答情報が複数ある場合に、ユーザが前提条件として使用した単語を元に基準を定め、分類して該回答情報を提示することを特徴とする請求項5または6に記載の情報検索方法。

【請求項8】 コンピュータに、ユーザからの入力文を解析して要素に分解する形態素解析ステップと、前記形態素解析ステップの結果に基づき、前記入力文を内部表現形式に変換して該入力文の概念および概念関係を明確にする意味解析ステップと、前記入力文における文と文の間の関係を抽出する文関係抽出ステップと、前記内部表現形式に基づき、データベースをアクセスして情報検索および情報登録を行うデータベース検索ステップと、前記情報検索の結果である回答情報の重複部分を抽出する回答情報検査ステップと、前記回答情報の重複部分を調整し提示順序の整合をとる回答情報整合ステップと、前記調整および整合後の回答情報を自然言語形式に変換する自然言語生成ステップと、を実行させるためのプログラムを記録した記録媒体。

【請求項9】 前記回答情報検査ステップにおける調整および整合は、前記意味解析ステップによる文概念および概念関係情報、並びに、前記文関係抽出ステップによる文関係情報に基づいて行われることを特徴とする請求項8に記載の記録媒体。

【請求項10】 前記回答情報検査ステップにおける調整および整合は、前記回答情報が複数ある場合に、ユーザが前提条件として使用した単語を元に基準を定め、分類して該回答情報を提示することを特徴とする請求項8または9に記載の記録媒体。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】 本発明は、自然言語を介した情報検索システム、情報検索方法および該情報検索方法を実行させるためのプログラムを記録した記録媒体に係り、特に、ユーザの検索要求における複数の意図や複雑な条件に対して情報の整理・分類を行って、検索要求に即した情報をユーザにわかりやすく提示する情報検索システム、情報検索方法および記録媒体に関する。

## 【0002】

【従来の技術】人間対人間の会話において、人間は、ある文を言いかけた直後に補足的な情報を付け加えたり、その前提条件を言い直したりすることがよくある。すなわち、人間は決して文法的に正しい文を1回の発話で行つてはいない。

【0003】一方、人間（ユーザ）と情報検索システム（コンピュータ）との自然言語を媒介としたインターフェースにおいては、1回のユーザの入力で完結した文が入力されることを前提としており、ユーザ側では、コンピュータからの回答を見て、必要と感じたら再度条件等の絞り込みなどを行う必要があった。

【0004】また、コンピュータ側が提示する文については、内容の冗長性の問題が指摘されている。初期の頃の情報検索システムにおいては、音声認識技術の稚拙さから、コンピュータは質問者（ユーザ）の質問をインターフェース上で繰り返した上で回答を提示し、確認のために冗長な情報のやり取りが行われていた。これに対して、1度質問や回答に使用された情報は繰り返して出力しないシステムも提案されている。例えば、特開昭62-169220号公報の「マンマシン・インターフェース方式」においては、人間とコンピュータの発話内容を記憶しておき、同じ情報を繰り返して提示しないようにして、人間とコンピュータ間の自然な会話の可能なシステムを目指している。

#### 【0005】

【発明が解決しようとする課題】しかしながら、上記従来の情報検索システムにおいても、情報を整理してユーザに提示するという機能は備えておらず、複数の文や複雑な条件が含まれる文に対しては整理された回答が生成されないという問題があった。

【0006】例えば、質問者（ユーザ）が「体温計で安いものを教えて下さい。できればメーカーはA社の体温計をお願いします。」という質問をコンピュータに対して発した場合に、従来、コンピュータ側では、これら2つの文の関連がわからずに、「体温計で安いものについてお答えします。△（B社），○○（A社），…です。」といった回答を提示するといったように、ユーザの複数の意図に即した情報がわかりやすく提示されていないという問題点である。

【0007】本発明は、このような従来の問題点に着目してなされたもので、その目的とするところは、ユーザの検索要求における複数の意図や複雑な条件に対して情報の整理・分類を行って、検索要求に即した情報をユーザにわかりやすく提示し、冗長な情報のやり取りを抑制して通信時間や通信コストを減少させた情報検索システム、情報検索方法および記録媒体を提供することにある。

#### 【0008】

【課題を解決するための手段】上記課題を解決するため、本出願の請求項1に記載の発明は、ユーザからの入

力文を解析して要素に分解する形態素解析手段と、前記形態素解析結果に基づき、前記入力文を内部表現形式に変換して該入力文の概念および概念関係を明確にする意味解析手段と、前記入力文における文と文の間の関係を抽出する文関係抽出手段と、前記内部表現形式に基づき、データベースをアクセスして情報検索および情報登録を行うデータベース検索手段と、前記情報検索の結果である回答情報の重複部分を抽出する回答情報検査手段と、前記回答情報の重複部分を調整し提示順序の整合をとる回答情報整合手段と、前記調整および整合後の回答情報を自然言語形式に変換する自然言語文生成手段とを具備することを特徴とする情報検索システムにある。

【0009】また、本出願の請求項2に記載の発明は、前記回答情報検査手段における調整および整合は、前記意味解析手段による文概念および概念関係情報、並びに、前記文関係抽出手段による文関係情報に基づいて行われることを特徴とする請求項1に記載の情報検索システムにある。

【0010】また、本出願の請求項3に記載の発明は、前記意味解析手段による前記内部表現形式、文概念および概念関係情報、並びに、前記文関係抽出手段による文関係情報を一時的に保持する質問バッファと、前記データベース検索手段による回答情報を一時的に保持する回答バッファとを具備することを特徴とする請求項1または2に記載の情報検索システムにある。

【0011】また、本出願の請求項4に記載の発明は、前記回答情報検査手段における調整および整合は、前記回答情報が複数ある場合に、ユーザが前提条件として使用した単語を元に基準を定め、分類して該回答情報を提示することを特徴とする請求項1、2または3に記載の情報検索システムにある。

【0012】また、本出願の請求項5に記載の発明は、ユーザからの入力文を解析して要素に分解する形態素解析ステップと、前記形態素解析ステップの結果に基づき、前記入力文を内部表現形式に変換して該入力文の概念および概念関係を明確にする意味解析ステップと、前記入力文における文と文の間の関係を抽出する文関係抽出ステップと、前記内部表現形式に基づき、データベースをアクセスして情報検索および情報登録を行うデータベース検索ステップと、前記情報検索の結果である回答情報の重複部分を抽出する回答情報検査ステップと、前記回答情報の重複部分を調整し提示順序の整合をとる回答情報整合ステップと、前記調整および整合後の回答情報を自然言語形式に変換する自然言語文生成ステップとを具備することを特徴とする情報検索方法にある。

【0013】また、本出願の請求項6に記載の発明は、前記回答情報検査ステップにおける調整および整合は、前記意味解析ステップによる文概念および概念関係情報、並びに、前記文関係抽出ステップによる文関係情報に基づいて行われることを特徴とする請求項5に記載の

情報検索方法にある。

【0014】また、本出願の請求項7に記載の発明は、前記回答情報検査ステップにおける調整および整合は、前記回答情報が複数ある場合に、ユーザが前提条件として使用した単語を元に基準を定め、分類して該回答情報を提示することを特徴とする請求項5または6に記載の情報検索方法にある。

【0015】また、本出願の請求項8に記載の発明は、コンピュータに、ユーザからの入力文を解析して要素に分解する形態素解析ステップと、前記形態素解析ステップの結果に基づき、前記入力文を内部表現形式に変換して該入力文の概念および概念関係を明確にする意味解析ステップと、前記入力文における文と文の間の関係を抽出する文関係抽出ステップと、前記内部表現形式に基づき、データベースをアクセスして情報検索および情報登録を行うデータベース検索ステップと、前記情報検索の結果である回答情報の重複部分を抽出する回答情報検査ステップと、前記回答情報の重複部分を調整し提示順序の整合をとる回答情報整合ステップと、前記調整および整合後の回答情報を自然言語形式に変換する自然言語文生成ステップとを実行させるためのプログラムを記録した記録媒体にある。

【0016】また、本出願の請求項9に記載の発明は、前記回答情報検査ステップにおける調整および整合は、前記意味解析ステップによる文概念および概念関係情報、並びに、前記文関係抽出ステップによる文関係情報に基づいて行われることを特徴とする請求項8に記載の記録媒体にある。

【0017】さらに、本出願の請求項10に記載の発明は、前記回答情報検査ステップにおける調整および整合は、前記回答情報が複数ある場合に、ユーザが前提条件として使用した単語を元に基準を定め、分類して該回答情報を提示することを特徴とする請求項8または9に記載の記録媒体にある。

【0018】そして、この請求項1、請求項5または請求項8に記載の発明によれば、形態素解析手段（形態素解析ステップ）によりユーザからの入力文を解析して要素に分解し、形態素解析手段（形態素解析ステップ）の結果に基づき、意味解析手段（意味解析ステップ）により入力文を内部表現形式に変換して該入力文の概念および概念関係を明確にし、文関係抽出手段（文関係抽出ステップ）により入力文における文と文の間の関係を抽出し、データベース検索手段（データベース検索ステップ）により内部表現形式に基づきデータベースをアクセスして情報検索および情報登録を行い、回答情報検査手段（回答情報検査ステップ）により情報検索の結果である回答情報の重複部分を抽出し、回答情報整合手段（回答情報整合ステップ）により回答情報の重複部分を調整し提示順序の整合をとり、自然言語文生成手段（自然言語文生成ステップ）により調整および整合後の回答情報

を自然言語形式に変換する。ここで、「内部表現形式」としては例えばフレーム形式や意味ネットワーク形式等が適用される。また、「入力文の概念および概念関係の明確化」には、情報検索の質問文かあるいは情報登録の情報提供文か等の発話タイプの同定や格解析による抽出結果の内部表現形式への展開などの処理が該当する。さらに、「文と文との関係」は、接続詞や副詞等から文関係情報（包含、並列、論理積、論理和、論理否定等）が抽出されて内部表現形式に反映される。

10 【0019】このように、ユーザの入力文（検索要求）における複数の質問等の意図や複雑な条件に対する回答情報について、重複部分を調整（除去等）し、また提示順序の整合をとって、検索要求に即した情報をユーザにわかりやすく提示するので、冗長な情報のやり取りを抑制して通信時間や通信コストを減少させることができる。

【0020】また、請求項2、請求項6または請求項9に記載の発明によれば、回答情報検査手段（回答情報検査ステップ）における調整および整合の処理を、意味解析手段（意味解析ステップ）による文概念および概念関係情報、並びに、文関係抽出手段（文関係抽出ステップ）による文関係情報に基づいて行う。「調整および整合の処理」は、例えば、包含、並列、論理積、論理和、論理否定等の文関係情報に基づいて、回答情報中の重複部分や同一の注目部分等を削除したり、回答情報の並べ替えを行ったりする。これにより、ユーザの入力文（検索要求）における複数の質問等の意図や複雑な条件に対する回答を行う際に、冗長な情報のやり取りを抑制して通信時間や通信コストを減少させることができる。

20 【0021】また、請求項3に記載の発明によれば、質問パッファに、意味解析手段による内部表現形式、文概念および概念関係情報、並びに、文関係抽出手段による文関係情報を一時的に保持し、また回答パッファにはデータベース検索手段による回答情報を一時的に保持する。これにより、複数の質問等の意図や複雑な条件を、内部表現形式、文概念および概念関係情報並びに文関係情報として質問パッファに一時的に保持でき、また、該質問に対する回答情報についても回答パッファに一時的に保持でき、一連の処理を円滑に行うことが可能となる。

40 【0022】さらに、請求項4、請求項7または請求項10に記載の発明によれば、回答情報検査手段（回答情報検査ステップ）における調整および整合の処理において、回答情報が複数ある場合に、ユーザが前提条件として使用した単語を元に基準を定め、分類して該回答情報を提示する。このように、複数の関連する回答が所定の基準に沿って分類・整理（並べ替え等）されて提示されるので、ユーザにとって理解しやすく、冗長な情報のやり取りを抑制して通信時間や通信コストを減少させることができる。

## 【0023】

【発明の実施の形態】以下、本発明の情報検索システム、情報検索方法および記録媒体の実施の形態について、図面を参照して詳細に説明する。

【0024】図1は本発明の一実施形態に係る情報検索システムの構成図である。同図において、本実施形態の情報検索システムは、入力部101、形態素解析部102、意味解析部103、文関係抽出部104、データベース検索部105、回答情報検査部106、回答情報整合部107、自然言語文生成部108、出力部109、データベース111、質問バッファ112および回答バッファ113を備えて構成されている。なお、データベース111には、本実施形態の情報検索システムの対象領域の知識、情報等が格納されているものとする。

【0025】また、図2には、本実施形態の情報検索システムにおける情報検索方法を説明するフローチャートを示す。まず図2のフローチャートを参照しながら、上記各構成要素の機能の概要と情報検索方法の全体的な処理の流れの概要を説明し、その後、各構成要素における処理の詳細について具体例を例示しながら説明していく。なお、本発明の記録媒体の実施の形態についてもこの情報処理検索方法の説明に含まれるものである。

【0026】本実施形態の情報検索方法の大まかな手順は、まずステップ201で、入力部101を介してユーザからの自然言語による入力文が入力されると、ステップ202（文入力処理）で、該自然言語文の形態素、構文、意味解析が行われてデータベース111に対する操作命令が生成された後、ステップ203（文生成処理）で、データベース111に対するアクセスの結果である回答情報が検査・整合されて自然言語による回答文が生成されて、ステップ204で出力部109を介して出力される。

【0027】ステップ202（文入力処理）では、形態素解析部102により入力文（自然言語文）を解析して要素（形態素）に分解するステップ211（形態素解析ステップ）と、形態素解析結果に基づき、意味解析部103により入力文を内部表現形式に変換して該入力文の概念および概念関係を明確にするべく発話タイプを同定するステップ212および意味解析を行うステップ213（意味解析ステップ）と、文関係抽出部104により入力文における文と文の間の関係を抽出・検査するステップ214（文関係抽出ステップ）と、データベース検索部105により内部表現形式に基づきデータベース111をアクセスして情報検索および情報登録を行うためのデータベース操作命令を生成するステップ215との処理がシーケンシャルに進められている。

【0028】また、ステップ203（文生成処理）では、データベース検索部105により操作命令を介してデータベース111にアクセスしてアクセス結果を得るステップ216（データベース検索ステップ）と、回答

情報検査部106により情報検査の結果である回答情報の重複部分を検査・抽出するステップ217（回答情報検査ステップ）と、回答情報整合部107により回答情報の重複部分を調整し提示順序の整合をとるステップ218（回答情報整合ステップ）と、自然言語文生成部108により調整および整合後の回答情報を自然言語形式に変換するステップ219（自然言語文生成ステップ）との処理がシーケンシャルに進められている。

【0029】次に、各構成要素における処理の詳細について、具体例を例示しながら図3から図13の説明図を参照して説明する。

【0030】まず、入力部101を介して、ユーザから自然言語による入力文が入力される。入力部101の実現形態は自然言語によって表現されている文章（音声・電子テキスト等）を入力する装置（例えば、音声認識装置、文字認識装置、キーボード、メールシステム）であればよく、特に限定されるものではない。また、アプリケーションによって選択可能であることが望ましい。また、入力データの種類によって入力方法を選択・変更できるようにしても良い。例えば、数字の羅列はキーボードを介し、他の文字列は音声入力手段を介するといった如くである。

【0031】入力された文は、形態素解析部102によって語および活用形を含む品詞情報等の形態素に分解される。形態素解析の実行例を図3に例示する。図3

(a) の第1入力文例に対して、図3 (b) の (イ) から (ヘ) に示すような形態素が抽出される。なお、形態素解析は周知の技術をそのまま利用すればよく、形態素解析ツールとしては、例えば、オムロン社製の「SuperMorpho Ver 1.62」や京都大学および奈良先端科学技術大学院大学によって開発された「JUMAN」等がある。

【0032】次に、意味解析部103により構文解析が行われ、入力文が情報検索の質問文であるか、或いは情報登録の情報提供文であるか等の発話タイプが同定される。この発話タイプの同定は、文末の文字列のパターン照合によって行われる。図4に、発話タイプの同定を行うためのパターンルールを例示する。図4 (a) ~

(c) に示すように、「～ですか？」、「～について教えて下さい。」、「～を知りたい。」といった文末文字列を含む文は質問文と同定され、また図4 (d) ,

(e) に示すように、「～です。」、「～した。」といった文末文字列を含む文は情報提供文であると同定される。なお、ユーザの入力文の発話の意図タイプを同定する処理についてより精度を高めるには、例えば、文献「熊本、伊藤、海老名；”支援対話におけるユーザ発話意図の認識－ユーザ発話文の解析に基づく統計的アプローチ－”電子情報通信学会論文誌D-II, Vol. J77, No. 5, pp1114-1123(1994).」で提案されている手法等を使用することも可能である。

【0033】また、意味解析部103では、形態素解析

結果に基づき、入力文を内部表現形式に変換して該入力文の概念および概念関係、すなわち各要素（名詞、名詞句、動詞等）間の意味的関連性を明確にする。ここでは、格解析による抽出結果の内部表現形式への展開処理について図5の説明図を参照して説明する。なお、意味解析については、文献「長尾真、佐藤理史、黒橋慎夫；岩波講座ソフトウェア科学15 自然言語処理」、岩波書店(1996)、第5章」に詳しい。

【0034】例えば、図5(a)に示す第2入力文例に対して、図5(b)に示すような動詞「発売する」の動詞格フレームを適用する。その適用結果は図5(c)に示す如くであり、「動作主格(は／が)」、「対象格(を)」、「目標格(に)」の各スロットに情報が埋められる。なお、図5に示した具体例では、表層格と深層格を組み合わせたものを例示したが、助詞が付属していくとも深層格情報のみで格フレームを埋めることも可能である。

【0035】このようにして抽出された格解析結果は内部表現形式に展開される。図6には内部表現形式の具体例を例示する。内部表現形式としては、図6(a)に示すようなフレーム形式や、図6(b)に示すような意味ネットワーク形式などが使用される。図6(a)および(b)は図5(a)の第2入力文例について展開された内部表現形式であり、格解析結果と共に発話タイプの情報（この具体例では「情報提供」）も付加されて質問バッファ112内に一時的に保持される。

【0036】発話タイプが質問文で、WH型疑問文（Whyを除くWhen, Where, What型疑問文）の場合は、該当するスロット値にマーク「?」が埋め込まれる。また、真偽疑問文（肯定か否定の答えが期待される疑問文）の場合は、疑問の対象にマーク「?」を付加するなどして区別される。なお、マーク「?」を付加する以外にもスロット毎にフラグ変数を用意する方法も可能である。すなわち、事実（前提条件として明らかになっているもの）と質問点（不明な点）とが区別されるものであれば何れの手法を用いても良い。

【0037】次に、文関係抽出部104により、複数入力される文と文の間の関係を調べる。この文関係情報の抽出は、文頭から接続詞や副詞等を取り出してパターンマッチングにより行われ、文と文との関係が包含関係、並列関係、論理積、論理和または論理否定等の論理関係にあるかが抽出される。つまり、文中の「特に」、「例えば」、「etc.」等を取り出して表層的関係を抽出するものである。

【0038】図7に文関係情報の抽出処理を説明するフローチャートを示す。まずステップ701～704では、文頭から接続詞や副詞等を順次取り出してパターンマッチングを行い、マッチする場合には（ステップ703で成功した場合には）、該当するリンク種別の評価を上げる。ここで、リンク種別は包含関係、並列関係、論

理積、論理和または論理否定等の論理関係を表すもので、具体的には「AND：論理積の並列関係」、「OR：論理和の並列関係」、「ESPECIALLY：包含関係」等で表記される。また、このリンク種別は、内部表現形式においてフレーム（またはネットワーク）がリンクされる場合に、フレーム間でどのような関連（リンク）を持つかを表す情報となる。

【0039】ステップ701で全てのパターンマッチングが終了すると、ステップ705に進んで、リンク種別の中でも最も評価の高いものを選択して（この評価値を $\alpha$ とし）、ステップ706で予め設定されているしきい値 $Th\alpha$ と最高評価値 $\alpha$ とを比較して、ステップ707では、該しきい値 $Th\alpha$ を越える最高評価値を持つリンク種別について、これを内部表現形式に付加する。

【0040】例えば、図6(a)のように、フレーム形式の内部表現形式においては、スロット「関連フレーム」に関連するフレームのポインタとリンク種別を挿入する。図6(a)の具体例では、スロット「関連フレーム」に、関連フレームポインタ「P-00009」とリンク種別「AND」が挿入されている。

【0041】また、他の具体例を示せば、「(1) Aについて教えて下さい。(2) できれば／特にBについて…」という入力文(1)および(2)については、AとBとが関連しあっていてBがAに包含される包含関係にあるとして、文(1)の内部表現形式におけるスロット「関連フレーム」には「00xxx/ESPECIALLY」（00xxxは文(2)の内部表現形式フレームへのポインタ）という情報が挿入される。

【0042】さらに同様にして、「(3) Aについて教えて下さい。(4) もしくは／それとも／またはBについて…」という入力文(3)および(4)については、AとBとが関連しあっていてAとBが論理和の並列関係にあるとして、文(3)の内部表現形式におけるスロット「関連フレーム」には「00xyz/OR」（00xyzは文(4)の内部表現形式フレームへのポインタ）という情報が挿入される。

【0043】次に、データベース検索部105では、内部表現形式に基づきデータベース111をアクセスして情報検索および情報登録を行うためのデータベース操作命令を生成する。より具体的には、文関係抽出部104で抽出された文関係から文の間およびスロット間の関係を決定してデータベース操作命令を発行する。ここでは、第3入力文例「(5)〇年に発売の体温計について教えて下さい。(6) できればA社のを教えて下さい。」が入力された場合を具体例として説明する。この第3入力文例について生成される内部表現形式は、フレーム形式を使用する場合、図8に示すようなものとなる。

【0044】すなわち、入力文(5)に対して製品フレーム(No=00013)が、入力文(6)に対して製品フレーム(No=00014)がそれぞれ生成される。入力文(5)および(6)と共に発話タイプは質問文であり、入力文(5)では製品の

種類が「体温計」で、発売日が「〇年」と特定され、製品名および値段について質問しているものとされ、また入力文(6)では発売元が「A社」に特定されて製品名について質問しているものとされている。また、入力文(5)および(6)の関係は包含関係にあるので、入力文(5)に対する製品フレーム (No=00013) のスロット「関連フレーム」には (00014/ESPECIALLY) の情報が付加されている。

【0045】データベース操作命令の生成は、質問バッファ112に格納されている内部表現形式をデータベース操作命令文に変換することによって行われる。ここでは、データベース操作言語としてSQLを使用した例を示す。

【0046】上記第3入力文例について、図8に示した内部表現形式から図9 (a) または (b) のデータベース操作命令が得られる。図9 (a) は2つの製品フレーム間のESPECIALLYリンクを加味せずに入力文(5)に対する製品フレーム (No=00013) のみから生成されたデータベース操作命令であり、図9 (b) は2つの製品フレーム間のESPECIALLYリンクを加味し、入力文(5)に対する製品フレーム (No=00013) および入力文(6)に対する製品フレーム (No=00014) から生成されたデータベース操作命令である。

【0047】以上、入力文が質問文である場合のデータベース操作命令の生成について例示したが、入力文が情報登録文である場合には、図5 (a) の第2入力文例に対して生成されるデータベース操作命令は、図9 (c) に示す如くなる。

【0048】また、データベース検索部105では、生成したデータベース操作命令をデータベース111に対して実行して情報登録または情報問合せ(検索)を行う。上記第3入力文例のように発話タイプが質問文の場合には、アクセス(検索)結果を内部表現形式に変換して回答バッファ113に一時的に格納する。

【0049】例えば、図9 (a) または (b) のデータベース操作命令を実行した結果、図8 (a) の製品フレーム (No=00013) を元に、図10に示すようなフレーム形式の内部表現形式が生成される。すなわち、製品フレーム (No=00033) において、発話タイプは回答文であり、疑問の対象となっていた製品名および値段のスロットにそれぞれ「C」および「xxx円」の情報が挿入され、また発売日のスロットについてはより情報量が増えて「〇月×日」の情報に書き換えられている。さらに、関連フレームのスロットには、元の情報(質問)との関連を示すべく情報(00013/ORIGINAL)が挿入されて、製品フレーム (No=00013) が元の質問である旨を表し、次の検索結果との関連を示すべく情報(00034/NEXT)が挿入されている。

【0050】次に、回答情報検査部106では、情報検索の結果である回答情報に重複部分がないか、また統合

できないかを検査する。具体的には、回答バッファ113内に保持されている回答情報について、データの重複があれば該当する情報に「削除マーク」を付加する。図11には、回答情報検査処理の手順を説明するフローチャートを示す。

【0051】まずステップ1101では、回答バッファ113の現在の回答情報と1つ前の回答情報との関連性(リンク種別)を検査する。そしてステップ1102ではリンク種別が「AND」かを判断して、「AND」の場合はステップ1103で、回答情報で重複するものを残し、それ以外のものに削除マークを付ける。またステップ1104ではリンク種別が「ESPECIALLY」かを判断して、「ESPECIALLY」の場合にはステップ1105で、回答情報で重複するものを残し、それ以外のものに保留マークを付ける。また、ステップ1106ではリンク種別が「OR」かを判断して、「OR」の場合にはステップ1107で、回答情報で重複するものを残し、それ以外のものを1つにまとめる。さらにステップ1108ではリンク種別が「NOT」かを判断して、「NOT」の場合にはステップ1109で、1つ前の回答情報から現在注目している回答情報と同じものに削除マークを付ける。

【0052】次に、回答情報整合部107では、回答情報の重複部分および統合可能部分を調整し提示順序の並べ替えを行って整合をとる。すなわち、「削除マーク」の付いた回答情報を削除し、続いて検索の結果である回答情報が複数あるような場合に該回答情報の並べ替えを行うものである。

【0053】並べ替え処理における基準は以下の通りである。すなわち、第1に、ユーザの質問文の中で強調された条件があれば、それを基準にする。例えば、内部表現形式がフレーム形式の場合には、質問バッファ112にある内部表現形式内のESPECIALLYリンクで参照されているスロットが該当し、意味ネットワーク形式の場合には、質問バッファ112にある内部表現形式内のESPECIALLYリンクで参照されている概念の上位概念が該当する。また第2に、検索条件となるスロット値によって回答情報を順序付けする。例えば、時間概念を持つスロットを優先させたり、スロット値が名詞の場合は辞書順としたりする。さらに第3に、同列の検索条件が複数ある場合には、スロット値が名詞の場合を優先させたり、スロット値のグループ数が少ないほうを基準とする。

【0054】図12には、回答情報整合処理の手順を説明するフローチャートを示す。同図を参照して、以下より詳細に説明する。まず、ステップ1201では、回答バッファ113で現在の回答情報(フレーム)はリンク種別を持つか否かを判断し、持たない場合には処理を終了し、持つ場合には以下の処理に移行する。

【0055】ステップ1202では、削除マークの付いている回答情報を回答バッファ113から取り除く。そしてステップ1203ではリンク種別が「AND」もしく

は「OR」かを判断して、「AND」もしくは「OR」の場合にはステップ1204で、回答情報を上記並べ替え基準に基づき検索条件によって整列する。またステップ1205ではリンク種別が「ESPECIALLY」かを判断して、「ESPECIALLY」の場合にはステップ1206で、無印の回答情報を上記並べ替え基準に基づき検索条件によって整列し、同様に、保留マークの付いた回答情報についても整列する。さらにステップ1207ではリンク種別が「NOT」かを判断して、「NOT」の場合にはステップ1208で、回答情報を上記並べ替え基準に基づき検索条件によって整列する。

【0056】次に、自然言語文生成部108では、調整および整合後の内部表現形式を自然言語形式に変換して回答文を生成する。この回答文の生成は、図13に示すようなフレーズ辞書を用いて行う。フレーズ辞書には文の出力パターンを備え、該出力パターンにおける要素の一部が空になっていて該空き部分を埋めるべき条件である内部概念および制約が付加されて構築されている辞書である。すなわち、内部概念と制約を満たすフレーズを選択し、出力パターンの空き部分を埋めて、態および時制を整えた後に回答文が生成される。

【0057】さらに、生成された自然言語による回答文は、出力部109を介して出力される。出力部109の実現形態は、例えば表示手段としてのディスプレイや、音声合成装置等がある。

【0058】以上説明したように、本実施形態の情報検索システムおよび情報検索方法では、形態素解析部102（形態素解析ステップ）によりユーザからの入力文を解析して要素に分解し、形態素解析部102（形態素解析ステップ）の結果に基づき、意味解析部103（意味解析ステップ）により入力文を内部表現形式に変換して該入力文の概念および概念関係を明確にし、文関係抽出部104（文関係抽出ステップ）により入力文における文と文の間の関係を抽出し、データベース検索部105（データベース検索ステップ）により内部表現形式に基づきデータベース111をアクセスして情報検索および情報登録を行い、回答情報検査部106（回答情報検査ステップ）により情報検索の結果である回答情報の重複部分を抽出し、回答情報整合部107（回答情報整合ステップ）により回答情報の重複部分を調整し提示順序の整合をとり、自然言語文生成部108（自然言語文生成ステップ）により調整および整合後の回答情報を自然言語形式に変換するようにしており、ユーザの入力文（検索要求）における複数の質問等の意図や複雑な条件に対する回答情報について、重複部分を調整（除去等）し、また提示順序の整合をとて、検索要求に即した情報をユーザにわかりやすく提示するので、冗長な情報のやり取りを抑制して通信時間や通信コストを減少させることができる。

【0059】特に、回答情報検査部106における調整

および整合の処理を、意味解析部103による文概念および概念関係情報、並びに、文関係抽出部104による文関係情報、例えば、包含、並列、論理積、論理和、論理否定等の文関係情報に基づいて、回答情報中の重複部分や同一の注目部分等を削除したり、回答情報の並べ替えにより行うので、ユーザの入力文（検索要求）における複数の質問等の意図や複雑な条件に対する回答を行う際に、冗長な情報のやり取りを抑制して通信時間や通信コストを減少させることができる。

10 【0060】さらに、回答情報検査部106における調整および整合の処理において、回答情報が複数ある場合に、ユーザが前提条件として使用した単語を元に基準を定め、分類して該回答情報を提示するようにし、複数の関連する回答が所定の基準に沿って分類・整理（並べ替え等）されて提示されるので、ユーザにとって理解しやすく、冗長な情報のやり取りを抑制して通信時間や通信コストを減少させることができる。

#### 【0061】

【発明の効果】以上の説明から明らかなように、本発明によれば、形態素解析手段（形態素解析ステップ）によりユーザからの入力文を解析して要素に分解し、形態素解析手段（形態素解析ステップ）の結果に基づき、意味解析手段（意味解析ステップ）により入力文を内部表現形式に変換して該入力文の概念および概念関係を明確にし、文関係抽出手段（文関係抽出ステップ）により入力文における文と文の間の関係を抽出し、データベース検索手段（データベース検索ステップ）により内部表現形式に基づきデータベースをアクセスして情報検索および情報登録を行い、回答情報検査手段（回答情報検査ステップ）により情報検索の結果である回答情報の重複部分を抽出し、回答情報整合手段（回答情報整合ステップ）により回答情報の重複部分を調整し提示順序の整合をとり、自然言語文生成手段（自然言語文生成ステップ）により調整および整合後の回答情報を自然言語形式に変換することとしたので、ユーザの入力文（検索要求）における複数の質問等の意図や複雑な条件に対する回答情報について、重複部分を調整（除去等）し、また提示順序の整合をとて、検索要求に即した情報をユーザにわかりやすく提示でき、冗長な情報のやり取りを抑制して通信時間や通信コストを減少させ得る情報検索システム、情報処理方法および記

40 【0062】また、本発明によれば、回答情報検査手段（回答情報検査ステップ）における調整および整合の処理を、意味解析手段（意味解析ステップ）による文概念および概念関係情報、並びに、文関係抽出手段（文関係抽出ステップ）による文関係情報に基づいて行うこととしたので、ユーザの入力文（検索要求）における複数の質問等の意図や複雑な条件に対する回答を行う際に、冗長な情報のやり取りを抑制して通信時間や通信コストを減少させ得る情報検索システム、情報処理方法および記

録媒体を提供することができる。

【0063】また、本発明によれば、質問パッファに、意味解析手段による内部表現形式、文概念および概念関係情報、並びに、文関係抽出手段による文関係情報を一時的に保持し、また回答パッファにはデータベース検索手段による回答情報を一時的に保持することとしたので、複数の質問等の意図や複雑な条件を、内部表現形式、文概念および概念関係情報並びに文関係情報として質問パッファに一時的に保持でき、また、該質問に対する回答情報についても回答パッファに一時的に保持でき、一連の処理を円滑に行うことが可能となる。

【0064】さらに、本発明によれば、回答情報検査手段(回答情報検査ステップ)における調整および整合の処理において、回答情報が複数ある場合に、ユーザが前提条件として使用した単語を元に基準を定め、分類して該回答情報を提示することとしたので、複数の関連する回答が所定の基準に沿って分類・整理(並べ替え等)されて提示されるので、ユーザにとって理解しやすく、冗長な情報のやり取りを抑制して通信時間や通信コストを減少させ得る情報検索システム、情報処理方法および記録媒体を提供することができる。

#### 【図面の簡単な説明】

【図1】本発明の一実施形態に係る情報検索システムの構成図である。

【図2】実施形態の情報検索システムにおける情報検索方法の概要を説明するフローチャートである。

【図3】形態素解析部における形態素解析の実行例を示す説明図である。

【図4】意味解析部において発話タイプの同定を行うためのパターンルールを例示する説明図である。

【図5】意味解析部において格解析による抽出結果の内部表現形式への展開処理について例示する説明図である。

【図6】内部表現形式を例示する説明図であり、図6(a)はフレーム形式、図6(b)は意味ネットワーク

【図3】

(a) 第1入力文例  
「昨日〇〇の温度計を買いました」

(b) 形態素解析の結果  
(イ) 昨日:名詞  
(ロ) 〇〇:有名詞(メーカー名)  
(ハ) の:助詞  
(ニ) 温度計:名詞  
(ホ) を:助詞  
(ヘ) 買:買い:いました。:ワ行五段:

発話タイプを同定する  
パターンルール例:  
(a) ~ですか? → 質問  
(b) ~について教えて下さい。 → 質問  
(c) ~を知りたい。 → 質問  
(d) ~です。 → 情報提供  
(e) ~した。 → 情報提供

【図4】

(a) 第2入力文例  
「A社は、温度センサBを〇月△日に発売する。」

(b) 動詞「発売する」の  
動詞格フレームの例:  
(発売する  
(動作主格(は/が) <会社>)  
(対象格(を) <製品>)  
(目標格(に) <時間>))

(c) (発売する  
(動作主格(は/が) A社)  
(対象格(を) 温度センサB)  
(目標格(に) 〇月△日))

【図5】

形式である。

【図7】文関係抽出部における文関係情報の抽出処理を説明するフローチャートである。

【図8】第3入力文例について生成される内部表現形式(フレーム形式)を例示する説明図である。

【図9】図9(a), (b)は図8の内部表現形式から得られるデータベース操作命令を例示する説明図であり、図9(a)はESPECIALLYリンクを加味しない場合、図9(b)はESPECIALLYリンクを加味した場合であり、

10 図9(c)は図5(a)の第2入力文例(情報登録文)から得られるデータベース操作命令を例示する説明図である。

【図10】図9(a)または(b)のデータベース操作命令実行結果として生成される内部表現形式を例示する説明図である。

【図11】回答情報検査部における回答情報検査処理の手順を説明するフローチャートである。

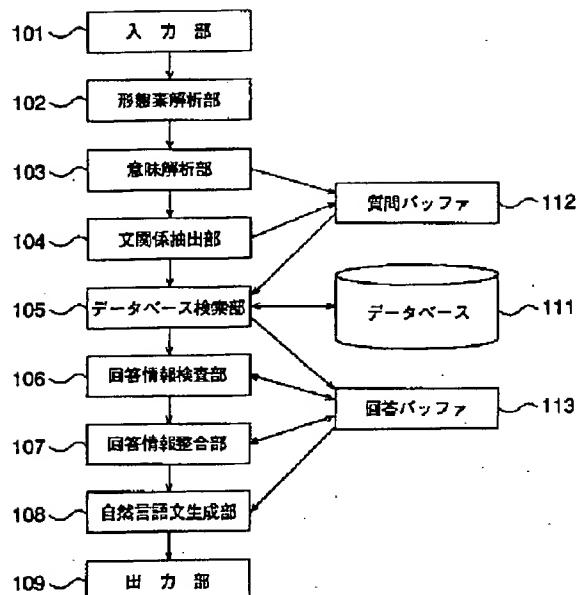
【図12】回答情報整合部における回答情報整合処理の手順を説明するフローチャートである。

20 【図13】自然言語文生成部において参照されるフレーズ辞書を例示する説明図である。

#### 【符号の説明】

101 入力部  
102 形態素解析部  
103 意味解析部  
104 文関係抽出部  
105 データベース検索部  
106 回答情報検査部  
107 回答情報整合部  
30 108 自然言語文生成部  
109 出力部  
111 データベース  
112 質問パッファ  
113 回答パッファ

【図1】



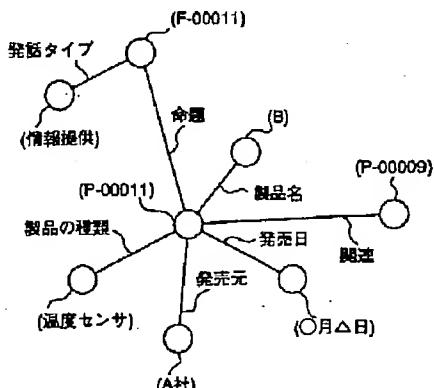
【図6】

(a) フレーム形式による内部表現

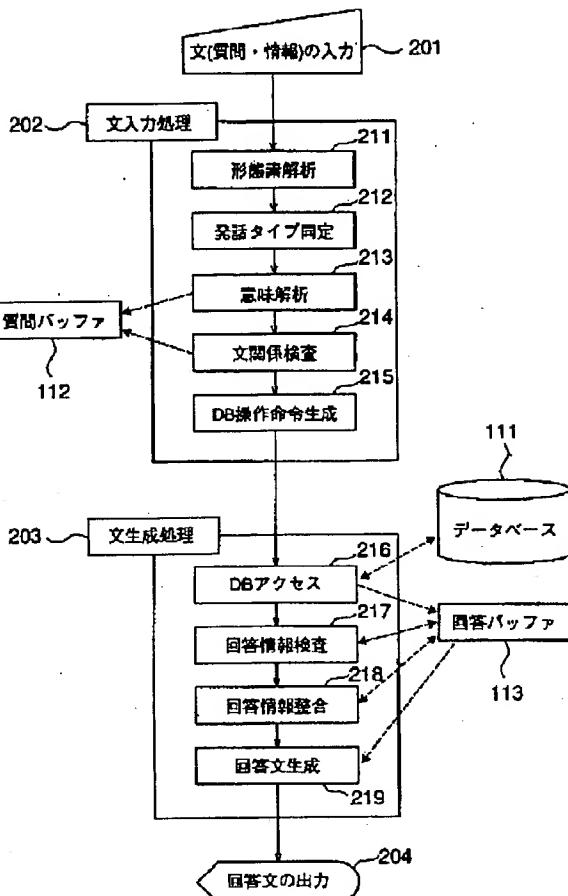
F-00011	
発話タイプ	情報提供
命題	P-00011, ...

P-00011	
製品名	B
製品の種類	温度センサ
発売元	A社
発売日	○月△日
...	
関連フレーム	(P-00009/AND)

(b) 意味ネットワーク形式による内部表現



【図2】



【図8】

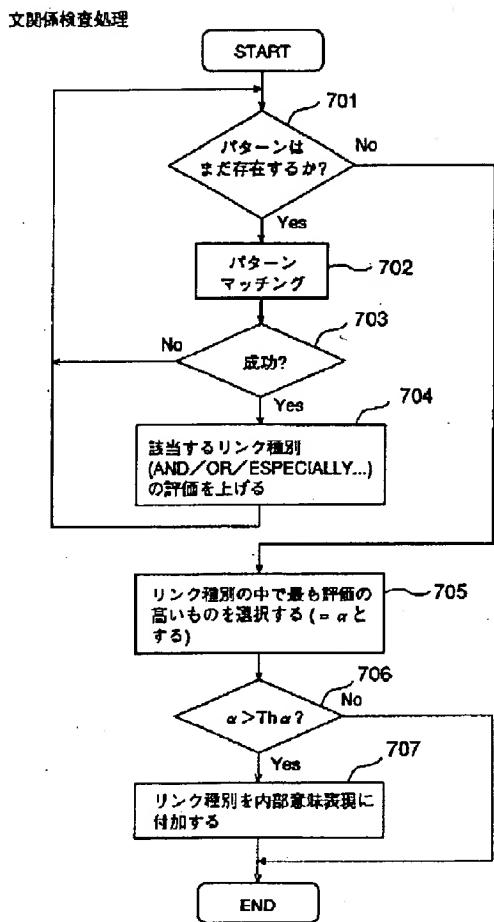
製品フレーム No=00013	
発話タイプ	質問
製品名	?
製品の種類	体温計
発売元	
発売日	○年
...	
値段	?
関連フレーム	(00014 / ESPECIALLY)

製品フレーム No=00033	
発話タイプ	回答
製品名	C
製品の種類	体温計
発売元	
発売日	○月×日
...	
値段	xxx円
関連フレーム	(00013/ORIGINAL) (00034/NEXT)

製品フレーム No=00014	
発話タイプ	質問
製品名	?
製品の種類	
発売元	A社
発売日	
...	
値段	
関連フレーム	

【図10】

【図7】



【図9】

(a) 命令文1 (ESPECIALLYリンクを加味しなかった場合)

```

    SELECT 製品名, 價格
    FROM 製品
    WHERE 製品の種類=体温計 AND 発売日.年=○
  
```

(b) 命令文2 (ESPECIALLYリンクを加味した場合)

```

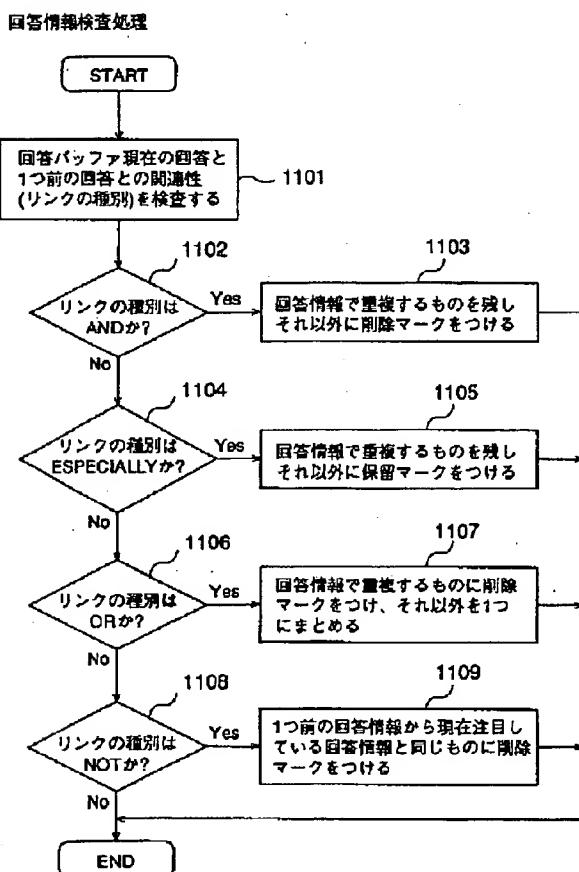
    SELECT 製品名, 價格
    FROM 製品
    WHERE 製品の種類=体温計 AND 発売元=A社
        AND 発売日=○月×日
  
```

(c) 命令文3 (情報登録命令)

```

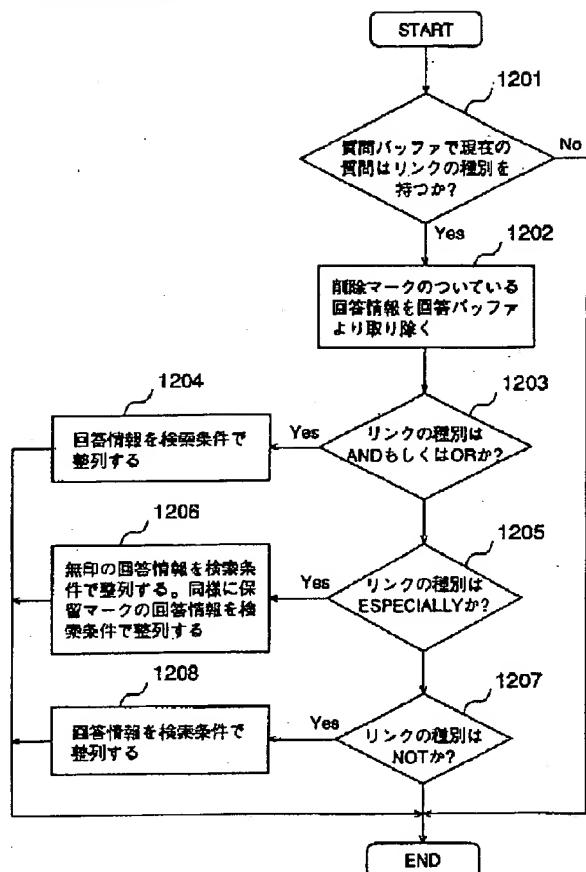
    INSERT INTO 製品 (発売元, 製品の種類, 製品名, 発売日)
    VALUES (A社, 体温計, B, ○月△日)
  
```

【図11】



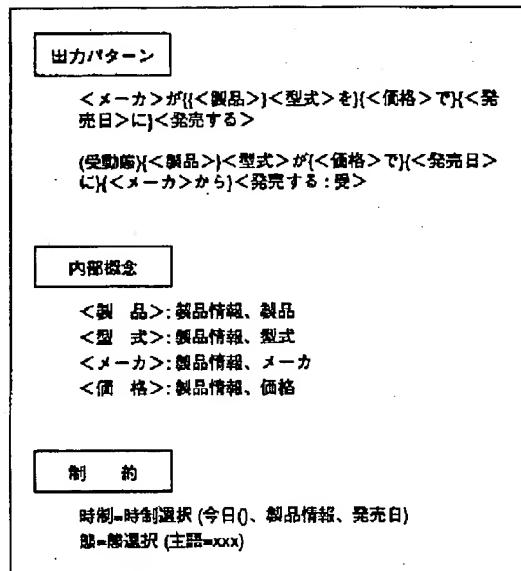
【図12】

## 回答情報整合処理



【図13】

## フレーズ(句語書)



フロントページの続き

(72)発明者 糸谷 和人

京都府京都市右京区花園土堂町10番地  
オムロン株式会社内